

Brownian bridges to submanifolds

James Thompson*

21st March 2017

Abstract

We introduce and study Brownian bridges to submanifolds. Our method involves proving a general formula for the integral over a submanifold of the minimal heat kernel on a complete Riemannian manifold. We use the formula to derive lower bounds, an asymptotic relation and derivative estimates. We also see a connection to hypersurface local time. This work is motivated by the desire to extend the analysis of path and loop spaces to measures on paths which terminate on a submanifold.

Keywords: Brownian bridge ; local time ; heat kernel ; submanifold

AMS MSC 2010: 58J65 ; 53B21 ; 60J55

Introduction

Brownian motion is a stochastic process naturally associated to any Riemannian manifold. The distance between Brownian motion and a submanifold was studied by the author in [46], using an inequality for the Laplacian of the distance function based on Jacobi field comparison. The Brownian bridge is given by conditioning Brownian motion to hit a fixed point at a fixed positive time. We extend this concept by replacing the fixed point with a submanifold. There has so far been very little research on this topic, even if the ambient space is Euclidean, although such processes have appeared in the context of Wiener measure approximation, as in [41].

Suppose that M is a complete and connected Riemannian manifold of dimension m and that N is a closed embedded submanifold of M of dimension $n \in \{0, \dots, m-1\}$. Assume for this introduction that M is compact. Denote by p^M the heat kernel on M and define

$$p_t^M(x, N) := \int_N p_t^M(x, y) d\text{vol}_N(y) \quad (1)$$

where vol_N denotes the measure on N induced by the embedding. For the case in which N is a point, vol_N is a delta measure and the integral reduces to a pointwise evaluation. The integrated heat kernel will be the central object of study. Some examples of it are given in Section 1. Following [46], it can be identified as the time derivative of a mean hypersurface local time. A comparison theorem for this object is included in Subsection 4.1.

In Section 2, denoting by $X(x)$ a Brownian motion on M starting at x , we fix $T > 0$ and prove that if $t \in [0, T)$ then for a bounded $\mathcal{F}_t^{X(x)}$ -measurable random variable F we have

$$\mathbb{E}[F | X_T(x) \in N] = \frac{\mathbb{E}[p_{T-t}^M(X_t(x), N)F]}{p_T^M(x, N)}.$$

*University of Luxembourg, Email: james.thompson@uni.lu

This gives rise to a diffusion over the time interval $[0, T)$ starting at x and arriving in N at time T , with time-dependent infinitesimal generator

$$\frac{1}{2}\Delta + \nabla \log p_{T-t}^M(\cdot, N).$$

We call this a *Brownian bridge to a submanifold*. To show that it is a semimartingale on $[0, T]$ we prove the gradient estimate

$$\|\nabla \log p_t^M(x, N)\|^2 \leq C \left(\frac{1}{t} + \frac{n}{t} \log \frac{1}{t} + \frac{d^2(x, N)}{t^2} \right)$$

and derive a Hessian estimate as corollary. These estimates are given in Section 6 by Theorem 6.2 and Corollary 6.4, respectively. We prove them using the method of Stroock [42], the inductive element of which had previously been discovered by Cheng, Li and Yau [7]. Our estimates generalize the main theorem of Engoulatov [14] and the gradient and Hessian estimates of Hsu [23], who considered only the one point case.

To prove the gradient estimate we require a suitable lower bound on $p_t^M(x, N)$. In Section 3 we define $r_N(\cdot) := d(\cdot, N)$ and introduce the diffusion on M starting at x with time dependent generator

$$\frac{1}{2}\Delta - \frac{r_N}{T-t} \frac{\partial}{\partial r_N}$$

where $\frac{\partial}{\partial r_N}$ denotes differentiation in the radial direction. We call such a process a *Fermi bridge*. It is an example of a *submanifold bridge process*, by which we mean a Brownian motion with drift which arrives in a submanifold at a fixed positive time. This terminology was introduced by Ndumu [30]. In Section 4 we prove the main result of this article, Theorem 4.1, which in terms of a Fermi bridge $\hat{X}(x)$ yields the formula

$$p_T^M(x, N) = (2\pi t)^{-\frac{(m-n)}{2}} \exp \left[-\frac{d^2(x, N)}{2t} \right] \lim_{t \uparrow T} \mathbb{E} \left[\exp \left[\int_0^t \frac{r_N(\hat{X}_s(x))}{T-s} (d\mathbf{A}_s + d\mathbf{L}_s) \right] \right].$$

Here $d\mathbf{A}$ is an absolutely continuous random measure, which takes into account the geometry of M in between N and the cut locus, while $d\mathbf{L}$ is a singular continuous random measure, which takes in to account the effect of the cut locus itself. If N is a point, this yields a formula for the heat kernel itself.

Theorem 4.1 is proved using a modification of the method of Elworthy and Truman [13]. Elworthy and Truman's heat kernel formula and Ndumu's extension of it to integrated heat kernels [30–33] typically require some invertibility of the exponential map and so cannot be applied directly to manifolds with a cut locus. The problem of extending these formulae to manifolds with a cut locus was posed to the author by Xue-Mei Li. Theorem 4.1 provides this extension and can also be adapted to include a drift and a potential, in analogy with the Schrödinger equation for a magnetic field, considered by Watling [50–52].

The formulae in [13], [30] and [50] are all expressed in terms of a stochastic process called the *semiclassical bridge*. The semiclassical bridge will be used to obtain derivative formulae for kernels of Schrödinger operators in a future publication by the author and Xue-Mei Li. It is similar to the Fermi bridge but has an additional drift which depends on the Jacobian determinant of the exponential map. This additional drift results in the appearance of a potential function which is difficult to understand geometrically. The random measure $d\mathbf{A}$ can, on the other hand, be written explicitly in terms of Jacobi fields.

This allows us in Section 5 to use Theorem 4.1 and the estimate on \mathbf{A} given in [46] to deduce a lower bound on the integrated heat kernel. This bound is stated in Theorem

5.1. It is used to prove an asymptotic relation, stated in Theorem 5.2, and in Section 6 it is used to prove the desired gradient and Hessian estimates. Potential applications of these results are discussed in Subsection 6.1.

Although for this introduction we have assumed that M is compact, this is not always necessary. In fact, we will usually only require some control on the Laplacian of the distance function. Geometric conditions for this were given in [46]. In particular, it was proved there that if there exists a function $\kappa : [0, \infty) \rightarrow [0, \infty)$ such that one of the following conditions is satisfied off the union of N and its cut locus:

- (C1) $n \in \{0, \dots, m-1\}$, the sectional curvature of planes containing the radial direction is bounded below by $-\kappa^2(r_N)$ and the absolute value of the principal curvature of N is bounded by a constant $\Lambda \geq 0$;
- (C2) $n = 0$ and the Ricci curvature in the radial direction is bounded below by $-(m-1)\kappa^2(r_N)$;
- (C3) $n = m-1$, the Ricci curvature in the radial direction is bounded below by $-(m-1)\kappa^2(r_N)$ and the absolute value of the mean curvature of N is bounded by a constant $\Lambda \geq 0$;

then the inequality

$$\frac{1}{2} \Delta r_N^2 \leq (m-n) + (n\Lambda + (m-1)\kappa(r_N)) r_N$$

holds off the cut locus pointwise and, by [40], on the whole of M in the sense of distributions. Conditions of this type are used to deduce Theorem 5.1.

The problem which originally motivated this piece of work was to extend Elworthy and Truman's formula to manifolds with a cut locus. Theorem 4.1 was the first new result to be proved in this article and led to the remainder of the investigation. The author wishes to thank Xue-Mei Li for for advising this and the study of point to submanifold bridges, as either conditioned Brownian motions or as Brownian motions with directed drift, and for suggesting several useful references.

1 The Heat Kernel Integrated over a Submanifold

Suppose that M is a complete and connected Riemannian manifold of dimension m and that $X(x)$ is a Brownian motion on M starting at $x \in M$ defined up to an explosion time $\zeta(x)$. An open connected subset D of M is called a *regular domain* if it has smooth boundary and compact closure. If p^D denotes the Dirichlet heat kernel on a regular domain D then p^D is the fundamental solution to the heat equation on D with Dirichlet boundary conditions. If f is a non-negative measurable function on D then p^D satisfies

$$\mathbb{E} [\mathbf{1}_{\{t < \tau_D\}} f(X_t(x))] = \int_M f(y) p_t^D(x, y) d \text{vol}_M(y)$$

for all $t > 0$ where τ_D denotes the first exit time of $X(x)$ from D . If $\{D_i\}_{i=1}^\infty$ is an exhaustion of M by regular domains then the minimal heat kernel on M is given by $p^M := \lim_{i \uparrow \infty} p^{D_i}$. It is the minimal fundamental solution of the heat equation on M and coincides with the transition densities of Brownian motion.

Example 1.1. On \mathbb{R}^m the heat kernel is given by the Gauss-Weierstrass kernel

$$p_t^{\mathbb{R}^m}(x, y) = (2\pi t)^{-\frac{m}{2}} \exp\left(-\frac{d^2(x, y)}{2t}\right) \quad (2)$$

for $x, y \in \mathbb{R}^m$ and $t > 0$.

Example 1.2. On \mathbb{H}_κ^3 , the hyperbolic space of dimension 3 with constant sectional curvatures $\kappa < 0$, there is the formula

$$p_t^{\mathbb{H}_\kappa^3}(x, y) = (2\pi t)^{-\frac{3}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right] \underbrace{\frac{\sqrt{-\kappa} d(x, y) e^{\frac{\kappa t}{2}}}{\sinh(\sqrt{-\kappa} d(x, y))}}_{\leq 1} \quad (3)$$

for $x, y \in \mathbb{H}_\kappa^3$ and $t > 0$.

Example 1.3. On \mathbb{S}^1 , the unit circle, there is the summation formula

$$p_t^{\mathbb{S}^1}(x, y) = (2\pi t)^{-\frac{1}{2}} \exp \left[-\frac{d^2(x, y)}{2t} \right] \underbrace{\sum_{k \in \mathbb{Z}} \exp \left[-\frac{2\pi k(d(x, y) + \pi k)}{t} \right]}_{\geq 1}$$

for $x, y \in \mathbb{S}^1$ and $t > 0$, a result of the fact that \mathbb{S}^1 is covered by \mathbb{R} .

The articles [18] and [29] provide iterative formulae for the heat kernels on the standard hyperbolic and spherical spaces, respectively, of arbitrary dimension. It is generally difficult to find deterministic formulae, unless the manifold in question exhibits a certain amount of symmetry, as in the above examples.

Now suppose that N is a closed embedded submanifold of M of dimension $n \in \{0, \dots, m-1\}$ and recall the definition of the integrated heat kernel given by formula (1). If D is a regular domain in M then define similarly

$$p_t^D(x, N) := \int_N p_t^D(x, y) d\text{vol}_N(y) \quad (4)$$

for $x \in M$ and $t > 0$, where vol_N denotes the measure on N induced by the embedding.

Example 1.4. For $r > 0$ denote by $\mathbb{S}^1(r)$ the circle of radius r centred at the origin in \mathbb{R}^2 . Then

$$p_t^{\mathbb{R}^2}(x, \mathbb{S}^1(r)) = r t^{-1} \exp \left[-\frac{(r^2 + \|x\|^2)}{2t} \right] \text{BesselI} \left(0, \frac{r\|x\|}{t} \right)$$

for $x \in \mathbb{R}^2$ and $t > 0$, where BesselI denotes the modified Bessel function of the first kind. A plot of the integrated kernel is given below in Figure 1.



Figure 1: A density plot of the integral over a circle of the heat kernel on \mathbb{R}^2 , at a fixed small time $t > 0$. The origin is located at the center of the image.

For $y \in M$ one can think of $p^M(\cdot, y)$ as a solution to the heat equation on M for the measure-valued initial condition given by the Dirac measure based at y . Similarly, $p^M(\cdot, N)$ can be thought of as a solution to the heat equation on M for the measure-valued initial condition vol_N . For example, if N is a closed embedded surface in \mathbb{R}^3 uniformly heated at time zero then $p^{\mathbb{R}^3}(\cdot, N)$ describes how the heat diffuses for positive times. Alternatively, if N is a closed embedded loop in \mathbb{R}^3 then $p^{\mathbb{R}^3}(\cdot, N)$ could be used to describe the temperature around a hot metal wire. A probabilistic interpretation of the integrated heat kernel will be given in the next section by Theorem 2.1.

Our basic approach will be to compare $p^M(\cdot, N)$ to the function

$$q_t(x, N) := (2\pi t)^{-\frac{(m-n)}{2}} \exp \left[-\frac{d^2(x, N)}{2t} \right] \quad (5)$$

since if \mathbb{R}^n is viewed as an affine linear subspace of \mathbb{R}^m then

$$p_T^{\mathbb{R}^m}(x, \mathbb{R}^n) = q_T(x, \mathbb{R}^n). \quad (6)$$

Example 1.5. Viewing the 2-dimensional hyperbolic space \mathbb{H}^2 as an embedded totally geodesic submanifold of \mathbb{H}^3 , Ndumu [30] used formula (3) with $\kappa = -1$ to show that

$$p_t^{\mathbb{H}^3}(x, \mathbb{H}^2) = q_t(x, \mathbb{H}^2) \frac{e^{-\frac{t}{2}}}{\cosh(d(x, \mathbb{H}^2))} \quad (7)$$

for $x \in \mathbb{H}^3$ and $t > 0$.

2 Brownian Bridges

Suppose in this section that M is stochastically complete and that N is compact. Fix $T > 0$ and $x \in M$ and consider the associated canonical probability space

$$(W(M), \mathcal{B}(W(M)), \mathbb{P}^x)$$

equipped with canonical filtration $\{\mathcal{B}_t(W(M))\}_{0 \leq t \leq T}$. Here $W(M)$ is the space of continuous paths defined on $[0, T]$ taking values in M , $\mathcal{B}_t(W(M))$ is the σ -algebra generated by the coordinate maps upto time t and \mathbb{P}^x is Wiener measure, with respect to which the coordinate process $\{X_t : t \in [0, T]\}$ is a Brownian motion on M starting at x .

2.1 Conditioning on the Distance Function

Recall that the *normal exponential map* is simply the exponential map restricted to the normal bundle TN^\perp . Denote its Jacobian determinant by θ_N . Inequalities for this object were provided in [46]. The following theorem shows how the integrated heat kernel appears naturally when a Brownian motion is conditioned to arrive in N at the fixed time T .

Theorem 2.1. Choose $t \in [0, T]$ and a bounded $\mathcal{B}_t(W(M))$ -measurable random variable F . Then

$$\mathbb{E}^x [F(X) | X_T \in N] = \frac{\mathbb{E}^x [p_{T-t}^M(X_t, N) F(X)]}{p_T^M(x, N)}. \quad (8)$$

Proof. For sufficiently small $\epsilon > 0$ it follows from the definition of conditional expectation, the Markov property, Fubini's theorem and the smooth coarea formula that

$$\begin{aligned}
& \mathbb{E}^x [F(X) | d(X_T, N) < \epsilon] \\
&= \frac{\mathbb{E}^x [\mathbf{1}_{\{d(X_T, N) < \epsilon\}} F(X)]}{\mathbb{P}^x \{d(X_T, N) < \epsilon\}} \\
&= \frac{\mathbb{E}^x [P_{T-t} \mathbf{1}_{\{d(\cdot, N) < \epsilon\}}(X_t) F(X)]}{\mathbb{P}^x \{d(X_T, N) < \epsilon\}} \\
&= \frac{\int_{B_\epsilon(N)} \mathbb{E}^x [p_{T-t}^M(X_t, y) F(X)] d\text{vol}_M(y)}{\int_{B_\epsilon(N)} p_T^M(x, y) d\text{vol}_M(y)} \\
&= \frac{\int_N \int_{B_\epsilon^z(0)} \mathbb{E}^x [p_{T-t}^M(X_t, \exp(\xi)) F(X)] \theta_N(\xi) d\xi d\text{vol}_N(z)}{\int_N \int_{B_\epsilon^z(0)} p_T^M(x, \exp(\xi)) \theta_N(\xi) d\xi d\text{vol}_N(z)}
\end{aligned}$$

where $B_\epsilon^z(0)$ denotes the open ball in $T_z N^\perp$ of radius ϵ centred at the origin. Since the volume of these balls is constant and independent of z , it follows by the continuity of the above integrands and $\theta_N|_N = 1$ that

$$\lim_{\epsilon \downarrow 0} \mathbb{E}^x [F(X) | d(X_T, N) < \epsilon] = \frac{\int_N \mathbb{E}^x [p_{T-t}^M(X_t, z) F(X)] d\text{vol}_N(z)}{\int_N p_T^M(x, z) d\text{vol}_N(z)}$$

from which the result follows, by the definition of the left-hand side of (8) as a Radon-Nikodym derivative. \square

For each $t \in [0, T)$ it follows, by Theorem 2.1 and the smooth coarea formula, that conditioning a Brownian motion to be in the *interior* of a tubular neighbourhood of N of radius ϵ at time T while separately conditioning Brownian motion to belong to the *boundary* of that tubular neighbourhood at time T results in two measures on $\mathcal{B}_t(W(M))$ which converge weakly to the same limit as $\epsilon \downarrow 0$.

2.2 Existence of the Bridge Measure

If for $y \in M$ we define a measure $\mathbb{P}^{x,y;T}$ on $\mathcal{B}(W(M))$ by $\mathbb{P}^{x,y;T}\{A\} = \mathbb{P}^x\{A | X_T = y\}$, for $A \in \mathcal{B}(W(M))$, then Theorem 2.1 implies that $\mathbb{P}^{x,y;T}$ is absolutely continuous with respect to \mathbb{P}^x on $\mathcal{B}_t(W(M))$ for any $t \in [0, T)$ and that the Radon-Nikodym derivative is given by

$$\frac{d\mathbb{P}^{x,y;T}|_{\mathcal{B}_t(W(M))}}{d\mathbb{P}^x} = \frac{p_{T-t}^M(X_t, y)}{p_T^M(x, y)}.$$

In particular, if $\mathbb{P}^{x,y;T}$ exists as a probability measure on the space of continuous paths starting at x and terminating at y at time T then under $\mathbb{P}^{x,y;T}$ and for $0 < t_1 < \dots < t_k < T$ the joint density function of X_{t_1}, \dots, X_{t_k} , denoted by $p_{t_1, \dots, t_k}^M(x, x_1, \dots, x_k, y)$, is given by

$$p_{t_1, \dots, t_k}^M(x, x_1, \dots, x_k, y) = \frac{p_{t_1}^M(x, x_1) p_{t_2-t_1}^M(x_1, x_2) \dots p_{T-t_k}^M(x_k, y)}{p_T^M(x, y)}, \quad (9)$$

as is well-known. We will prove the existence of $\mathbb{P}^{x,y;T}$ assuming that there exist constants $c, \sigma^2 > 0$ such that

$$p_t^M(w, z) \leq ct^{-\frac{m}{2}} \exp \left[-\frac{d^2(w, z)}{\sigma^2 t} \right] \quad (10)$$

for all $w, z \in M$ and $t \in (0, T]$ and a constant $\beta \geq 0$ such that

$$\theta_w(\xi) \leq \exp [\beta(1 + \|\xi\|^2)] \quad (11)$$

for all $w \in M$ and $\xi \in T_w M$. By [17] and [46], such bounds exist if the Ricci curvature is bounded below by a constant with vol_M lower regular. In particular, by [10] and [46], such bounds exist if the Ricci curvature is bounded with M having positive injectivity radius, in which case the constants c and σ^2 can be chosen so that σ^2 is arbitrarily close to 2. For the case in which M is compact, a version of the following lemma was proved by Driver in [11].

Lemma 2.2. *Assume (10) and (11), suppose $0 \leq s < t \leq T$ and without loss of generality assume $\beta > 0$. Then for all $\gamma \in (0, 1)$ there exists a constant $C(m, c, \sigma^2, \beta, \gamma, T) > 0$ such that for all $p > 0$ we have*

$$\mathbb{E}^{x, y; T} [d^p(X_s, X_t)] \leq \frac{C(m, c, \sigma^2, \beta, \gamma, T) \Gamma(\frac{m+p}{2})}{p_T^M(x, y) \Gamma(\frac{m}{2})} \left(\frac{\sigma^2(t-s)}{1-\gamma} \right)^{\frac{p}{2}}$$

so long as $t-s < \gamma(\sigma^2\beta)^{-1}$.

Proof. First assume $0 < s < t < 2T/3$. If $a > 0$ then

$$\int_0^\infty \exp[-ar^2] r^{m-1} ds = \frac{\Gamma(m/2)}{2a^{\frac{m}{2}}}$$

and so for $w \in M$ we have

$$\begin{aligned} & \int_M p_{t-s}^M(w, z) d^p(w, z) d\text{vol}_M(z) \\ & \leq c(t-s)^{-\frac{m}{2}} \int_M \exp\left[-\frac{d^2(w, z)}{\sigma^2(t-s)}\right] d^p(w, z) d\text{vol}_M(z) \\ & \leq ce^\beta(t-s)^{-\frac{m}{2}} \int_{T_w M} \|\xi\|^p \exp\left[\left(\beta - \frac{1}{\sigma^2(t-s)}\right) \|\xi\|^2\right] dv \\ & = ce^\beta \frac{m\pi^{\frac{m}{2}}(t-s)^{-\frac{m}{2}}}{\Gamma(\frac{m}{2}+1)} \int_0^\infty r^{p+m-1} \exp\left[\left(\beta - \frac{1}{\sigma^2(t-s)}\right) r^2\right] dr \\ & = ce^\beta \frac{\pi^{\frac{m}{2}} \Gamma(\frac{m+p}{2}) (t-s)^{-\frac{m}{2}}}{\Gamma(\frac{m}{2})} \left(\frac{1-\beta\sigma^2(t-s)}{\sigma^2(t-s)}\right)^{-\frac{(m+p)}{2}} \\ & \leq ce^\beta \frac{\Gamma(\frac{m+p}{2})}{\Gamma(\frac{m}{2})} \left(\frac{\sigma^2\pi}{1-\gamma}\right)^{\frac{m}{2}} \left(\frac{\sigma^2(t-s)}{1-\gamma}\right)^{\frac{p}{2}}. \end{aligned}$$

Thus there exists a constant $C_0(m, c, \sigma^2, \beta, \gamma) > 0$ such that

$$\int_M p_{t-s}^M(w, z) d^p(w, z) d\text{vol}_M(z) \leq C_0(m, c, \sigma^2, \beta, \gamma) \frac{\Gamma(\frac{m+p}{2})}{\Gamma(\frac{m}{2})} \left(\frac{\sigma^2(t-s)}{1-\gamma}\right)^{\frac{p}{2}}$$

for all $p > 0$, $w \in M$ and s, t satisfying $t-s < \gamma(\sigma^2\beta)^{-1}$. For such s, t we see that

$$\begin{aligned} & \mathbb{E}^{x, y; T} [d^p(X_s, X_t)] \\ & = \int_M \int_M \frac{p_s^M(x, w) p_{t-s}^M(w, z) d^p(w, z) p_{T-t}^M(z, y)}{p_T^M(x, y)} d\text{vol}_M(w) d\text{vol}_M(z) \\ & \leq (T/3)^{-m/2} \frac{C_1(m, c, \sigma^2, \beta, \gamma) \Gamma(\frac{m+p}{2})}{p_T^M(x, y) \Gamma(\frac{m}{2})} \left(\frac{\sigma^2(t-s)}{1-\gamma}\right)^{\frac{p}{2}}. \end{aligned}$$

The same result is obtained for $T/3 < s < t < T$. The cases $s = 0$ or $t = T$ can be treated similarly. \square

If F is a bounded $\mathcal{B}_t(W(M))$ -measurable function on $W(M)$ for some $t \in [0, T]$ then it follows from Theorem 2.1 that

$$\mathbb{E}^x [F(X)|X_T \in N] = \frac{\int_N p_T^M(x, y) \mathbb{E}^{x, y; T} [F(X)] d \text{vol}_N(y)}{p_T^M(x, N)} \quad (12)$$

which implies, by Lemma 2.2, that for all $p \geq 2$ there exists $\epsilon, C_\epsilon > 0$ such that

$$\mathbb{E}^x [d^p(X_s, X_t)|X_T \in N] \leq C_\epsilon(t-s)^{\frac{p}{2}}$$

for all $0 \leq s < t \leq T$ with $t-s < \epsilon$. It follows from Kolmogorov's continuity theorem, by covering the interval $[0, T]$ with finitely many closed intervals each of length less than ϵ , that there exists a probability measure $\mathbb{P}^{x; N, T}$ on the bridge space

$$L_{x; N, T}(M) := \{\omega \in W(M) : X_0(\omega) = x, X_T(\omega) \in N\}$$

which satisfies $\mathbb{P}^{x; N, T}\{A\} = \mathbb{P}^x\{A|X_T \in N\}$ for $A \in \mathcal{B}(W(M))$ and whose finite-dimensional distributions can be deduced from equations (9) and (12). In particular, if $\mathbb{P}_{X_T}^{x; N, T}$ denotes the law of the random variable X_T under the measure $\mathbb{P}^{x; N, T}$ then

$$\mathbb{P}_{X_T}^{x; N, T} = \frac{p_T^M(x, \cdot)}{p_T^M(x, N)} \text{vol}_N.$$

For example, if $M = \mathbb{R}^m$ with N the unit $(m-1)$ -sphere and $x = 0$ then $\mathbb{P}_{X_T}^{x; N, T}$ is given by the uniform measure on N . In contrast, if $M = \mathbb{R}^m$ with N an n -dimensional subspace and $x = 0$ then $\mathbb{P}_{X_T}^{x; N, T}$ is given by the heat kernel measure on N . Note that the measure $\mathbb{P}^{x; N, T}$ clearly exists in the latter case, even though N is non-compact.

Using Lemma 2.2 we can prove the following concentration inequality for tubular neighbourhoods.

Proposition 2.3. *Assuming (10) and (11), for all $\gamma \in (0, 1)$ there exists $\epsilon > 0$ such that*

$$\lim_{r \uparrow \infty} \frac{1}{r^2} \log \mathbb{P}^{x; N, T}\{X_t \notin B_r(N)\} \leq -\frac{1-\gamma}{\sigma^2(T-t)}$$

for all $t \in (T-\epsilon, T]$.

Proof. Without loss of generality, assume $\beta > 0$. Since $d(\cdot, N) = \inf_{y \in N} d(\cdot, y)$ it follows that for each $y \in N$ we have $\mathbb{E}^{x, y; T} [d^p(X_t, N)] \leq \mathbb{E}^{x, y; T} [d^p(X_t, y)]$. Therefore, by applying (12) with $F(X) = d^p(X_t, N)$, we see by Lemma 2.2 that for all $\gamma \in (0, 1)$ there exists $C(m, c, \sigma^2, \beta, \gamma, N, T) > 0$ such that

$$\mathbb{E}^{x; N, T} [d^p(X_t, N)] \leq \frac{C(m, c, \sigma^2, \beta, \gamma, N, T)}{p_T^M(x, N)} \frac{\Gamma(\frac{m+p}{2})}{\Gamma(\frac{m}{2})} \left(\frac{\sigma^2(T-t)}{1-\gamma} \right)^{\frac{p}{2}}$$

for all $0 \leq t \leq T$ with $T-t < \gamma(\sigma^2\beta)^{-1}$. For such t , choosing $\theta > 0$, applying this bound to the case where p is an even integer and summing yields

$$\mathbb{E}^{x; N, T} \left[e^{\frac{\theta}{2} d^2(X_t, N)} \right] \leq C(m, c, \sigma^2, \beta, \gamma, x, N, T) \left(1 - \frac{\theta \sigma^2(T-t)}{2(1-\gamma)} \right)^{-\frac{m}{2}}$$

so long as $t > T - 2(1-\gamma)(\theta\sigma^2)^{-1}$. Under these conditions on t , it follows from Markov's inequality that for all $r > 0$ there is the estimate

$$\mathbb{P}^{x; N, T}\{X_t \notin B_r(N)\} \leq C(m, \delta, \gamma, c_R, x, N, T) \left(1 - \frac{\theta \sigma^2(T-t)}{2(1-\gamma)} \right)^{-\frac{m}{2}} e^{-\frac{\theta r^2}{2}}.$$

Fixing $\delta \in [0, 1)$ and choosing $\theta = 2\delta(1-\gamma)(\sigma^2(T-t))^{-1}$ this yields

$$\lim_{r \uparrow \infty} \frac{1}{r^2} \log \mathbb{P}^{x; N, T}\{X_t \notin B_r(N)\} \leq -\frac{\delta(1-\gamma)}{\sigma^2(T-t)}$$

from which the result follows since δ can be chosen arbitrarily close to 1. \square

2.3 Semimartingale Property

It follows from (12) and Girsanov's theorem that under the measure $\mathbb{P}^{x;N,T}$ the coordinate process on $W(M)$ is a diffusion on the half-open time interval $[0, T)$ starting at x with time-dependent infinitesimal generator

$$\frac{1}{2}\Delta + \nabla \log p_{T-t}^M(\cdot, N)$$

for $t \in [0, T)$. To show that it is a semimartingale under this measure over the closed time interval $[0, T]$ we require a suitable estimate on the logarithmic derivative. This estimate will be deduced from the results obtained in Sections 3, 4 and 5. It is given by Theorem 6.2 in Section 6 and the semimartingale property is given by Corollary 6.4.

3 Fermi Bridges

In this section we introduce a bridge process defined in terms of the distance function $r_N(\cdot) := d(\cdot, N)$. Recall that the cut locus of N , denoted by $\text{Cut}(N)$, is given by the closure of the set of points at which r_N^2 fails to be differentiable. It is a closed set of vol_M -measure zero which is, upto a set of Hausdorff dimension at most $m - 2$, an at most countable union $\mathring{C}(N)$ of open subsets of hypersurfaces. These facts follow from the work of Angulo Ardoy and Guijarro [1] and Mantegazza and Menzucchi [27] who used Hamilton-Jacobi equations and the theory of viscosity solutions. The vector field $\frac{\partial}{\partial r_N}$ will denote differentiation in the radial direction, defined off the union of N and $\text{Cut}(N)$ to be the gradient of r_N and set equal to zero elsewhere. For $T > 0$ fixed with $q(\cdot, N)$ defined by equation (5), the time-dependent vector field

$$\mathbf{1}_{M \setminus \text{Cut}(M)} \nabla \log q_{T-t}(\cdot, N) = -\frac{r_N}{T-t} \frac{\partial}{\partial r_N}$$

where $t \in [0, T)$ is smooth away from the cut locus but generally not continuous on it. One imagines the deterministic flow associated to this vector field as being one for which $\text{Cut}(N)$ is a source and for which N is a sink. The strength of the flow increases dramatically as the terminal time T is approached, while the vector field vanishes on N . A diffusion on M starting at x with time-dependent infinitesimal generator

$$\frac{1}{2}\Delta - \frac{r_N}{T-t} \frac{\partial}{\partial r_N}$$

will be called a *Fermi bridge between x and N in time T* . Such processes will always be defined upto a predictable stopping time less than or equal to T . We use the name *Fermi bridge* since the time-dependent part of the drift acts in a direction normal to N , which would be the radial part of a system of polar Fermi coordinates, and since there are conditions under which this process arrives at N at time T almost surely. Note that if $M = \mathbb{R}^m$ with N a point then the above definition reduces to that of a standard Brownian bridge.

3.1 Radial Part

Suppose that $\hat{X}(x)$ is a Fermi bridge between x and N in time T , defined upto the minimum of T and its explosion time. Suppose also that D is a regular domain in M and denote by $\hat{\tau}_D$ the first exit time of $\hat{X}(x)$ from D . Since $\mathring{C}(N)$ is polar for $\hat{X}(x)$ and since the martingale part of an antidevelopment of $\hat{X}(x)$ is a standard Brownian motion, the Itô formula of Barden and Le [3, Theorem 1] implies that there exist continuous adapted

non-decreasing and non-negative processes $L^N(\hat{X}(x))$ and $L^{\hat{C}(N)}(\hat{X}(x))$ whose associated random measures are singular with respect to Lebesgue measure and supported when $\hat{X}(x)$ takes values in N and $\hat{C}(N)$, respectively, such that

$$\begin{aligned} r_N(\hat{X}_{t \wedge \hat{\tau}_D}(x)) &= r_N(x) + \beta_{t \wedge \hat{\tau}_D} + \frac{1}{2} \int_0^{t \wedge \hat{\tau}_D} \Delta r_N(\hat{X}_s(x)) ds \\ &\quad - \int_0^{t \wedge \hat{\tau}_D} \frac{r_N(\hat{X}_s(x))}{T-s} ds - \mathbb{L}_{t \wedge \hat{\tau}_D}^{\text{Cut}(N)}(\hat{X}(x)) + L_{t \wedge \hat{\tau}_D}^N(\hat{X}(x)) \end{aligned} \quad (13)$$

for all $t \in [0, T)$, almost surely, where β is a standard one-dimensional Brownian motion and

$$d\mathbb{L}^{\text{Cut}(N)}(\hat{X}(x)) := -\frac{1}{2} (D^+ - D^-)_{\hat{X}(x)} r_N(\mathbf{n}) dL^{\hat{C}(N)}(\hat{X}(x)). \quad (14)$$

Here \mathbf{n} is any unit normal vector field on $\hat{C}(N)$ and the Gâteaux derivatives $D^\pm r_N$ are defined for $z \in \hat{C}(N)$ and $v \in T_z M$ by

$$D_z^+ r_N(v) := \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} (f(\exp_z(\epsilon v)) - r_N(z))$$

and $D_z^- r_N(v) := -D_z^+ r_N(-v)$. The processes $L^N(\hat{X}(x))$ and $L^{\hat{C}(N)}(\hat{X}(x))$ are, roughly speaking, given by the local times at zero of $d(\hat{X}(x), N)$ and $d(\hat{X}(x), \hat{C}(N))$, respectively, whenever the latter make sense. Using this formula we can estimate the radial moments of the Fermi bridge, restricting our attention to the domain D and doing so under the assumption that there exist constants $\nu \geq 1$ and $\lambda \geq 0$ such that

$$\frac{1}{2} \Delta r_N^2 \leq \nu + \lambda r_N^2 \quad (15)$$

on $D \setminus \text{Cut}(N)$.

Theorem 3.1. *Let $\nu \geq 1$ and $\lambda \geq 0$ be any constants such that inequality (15) holds on $D \setminus \text{Cut}(N)$. Then we have*

$$\mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_D\}} r_N^2(\hat{X}_t(x)) \right] \leq \left(r_N^2(x) \left(\frac{T-t}{T} \right) + \nu t \right) \left(\frac{T-t}{T} \right) e^{\lambda t} \quad (16)$$

for all $t \in [0, T)$.

Proof. Define the function $\hat{f}_{x,2} : [0, T) \rightarrow \mathbb{R}$ by

$$\hat{f}_{x,2}(t) := \mathbb{E}[\mathbf{1}_{\{t < \hat{\tau}_D\}} r_N^2(\hat{X}_t(x))]$$

for $t \in [0, T)$. By Itô's formula and formula (13) we deduce the differential inequality

$$\begin{cases} \hat{f}'_{x,2}(t) \leq \nu + \left(\lambda - \frac{2}{T-t} \right) \hat{f}_{x,2}(t) \\ \hat{f}_{x,2}(0) = r_N^2(x) \end{cases}$$

for all $t \in [0, T)$. Applying Gronwall's inequality to it yields

$$\begin{aligned} \hat{f}_{x,2}(t) &\leq \left(r_N^2(x) + \nu \int_0^t \left(\frac{T}{T-s} \right)^2 e^{-s\lambda} ds \right) \left(\frac{T-t}{t} \right)^2 e^{\lambda t} \\ &\leq \left(r_N^2(x) + \nu t \left(\frac{t}{T-t} \right) \right) \left(\frac{T-t}{t} \right)^2 e^{\lambda t} \end{aligned}$$

where we used the assumption $\lambda \geq 0$ for the second inequality. \square

Note that for $M = \mathbb{R}^m$ with N a linear subspace, with $\nu = m - n$ and $\lambda = 0$, one can set $D = M$ and inequality (16) holds as an *equality*. Also, Jensen's inequality implies the following estimate on the first radial moment.

Corollary 3.2. *Let $\nu \geq 1$ and $\lambda \geq 0$ be any constants such that inequality (15) holds on $D \setminus \text{Cut}(N)$. Then we have*

$$\mathbb{E}[\mathbf{1}_{\{t < \hat{\tau}_D\}} r_N(\hat{X}_t(x))] \leq \left(\left(r_N^2(x) \left(\frac{T-t}{T} \right) + \nu t \right) \left(\frac{T-t}{T} \right) \right)^{\frac{1}{2}} e^{\frac{\lambda t}{2}}$$

for all $t \in [0, T)$.

In the next subsection we consider the case where there exist constants $\nu \geq 1$ and $\lambda \geq 0$ such that inequality (15) holds on the whole of $M \setminus \text{Cut}(N)$.

3.2 Bridge Property

Suppose for this subsection that there exists constants $\nu \geq 1$ and $\lambda \geq 0$ such that inequality (15) holds on $M \setminus \text{Cut}(N)$. Suppose also that $X(x)$ is a *non-explosive* Brownian motion on M starting at x , defined on a filtered probability space

$$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$$

satisfying the usual conditions. Note that, by [46, Theorem 5], if N is compact then inequality (15) *implies* the non-explosion of $X(x)$. For $t \in [0, T)$ define

$$M_t := \exp \left[- \int_0^t \frac{r_N(X_s(x))}{T-s} \left\langle \frac{\partial}{\partial r_N}, U_s dB_s \right\rangle - \frac{1}{2} \int_0^t \frac{r_N^2(X_s(x))}{(T-s)^2} ds \right]$$

where U is a horizontal lift of $X(x)$ to the orthonormal frame bundle with antidevelopment B . It follows from [46, Proposition 3] and Novikov's criterion that M is a martingale up to time t for each $t \in [0, T)$. For each $t \in [0, T)$ we can therefore define a new probability measure \mathbb{Q}_t on \mathcal{F}_t by $d\mathbb{Q}_t = M_t d\mathbb{P}$. It follows from Girsanov's theorem that the process $X(x)$ when restricted to $[0, t)$ and considered on the filtered probability space

$$(\Omega, \mathcal{F}_t, \{\mathcal{F}_s\}_{s \in [0, t)}, \mathbb{Q}_t)$$

is a Fermi bridge between x and N in time T . We therefore obtain a new process, denoted by $\hat{X}(x)$, defined on $[0, T)$ and equivalent to the Brownian motion $X(x)$ on $[0, t)$ for each $t \in [0, T)$. By considering an exhaustion of M by regular domains, Theorem 3.1 and the monotone convergence theorem imply that it satisfies the bridge property

$$\lim_{t \uparrow T} r_N(\hat{X}_t(x)) = 0$$

almost surely. In particular, if N is a point p then one then can extend $\hat{X}(x)$ to a continuous process on $[0, T]$ by setting $\hat{X}_T(x) = p$.

4 Heat Kernel Formula and Comparison Theorem

In this section we prove the main result of this article. Denote by $\mathcal{M}(N)$ the largest domain in TN^\perp with star-like fibres and such that $\exp|_{\mathcal{M}(N)}$ is a diffeomorphism onto its image. Then the image of $\exp|_{\mathcal{M}(N)}$ is $M \setminus \text{Cut}(N)$. Recalling that θ_N denotes the Jacobian determinant of the normal exponential, define also

$$\Theta_N := \theta_N \circ (\exp|_{\mathcal{M}(N)})^{-1}$$

and note that from [16] there is the formula

$$\frac{1}{2}\Delta r_N^2 = (m - n) + r_N \frac{\partial}{\partial r_N} \log \Theta_N \quad (17)$$

on $M \setminus \text{Cut}(N)$. Recall also the definition of the geometric local time $\mathbb{L}^{\text{Cut}(N)}(\hat{X}(x))$ given in Subsection 3.1 by formula (14).

Theorem 4.1. *Suppose that M is a complete and connected Riemannian manifold of dimension m , that N is a closed embedded submanifold of M of dimension $n \in \{0, \dots, m-1\}$ and that D is a regular domain in M . Suppose that $x \in M$ with $T > 0$ and that $\hat{X}(x)$ is a Fermi bridge between x and N in time T . Denote by $\hat{\tau}_D$ the first exit time of this process from D . Then, with $q_T(x, N)$ and $p_T^D(x, N)$ defined by (5) and (4), respectively, we have*

$$p_T^D(x, N) = q_T(x, N) \lim_{t \uparrow T} \mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_D\}} \exp \left[\int_0^t \frac{r_N(\hat{X}_s(x))}{T-s} (d\mathbf{A}_s + d\mathbf{L}_s) \right] \right]$$

where

$$d\mathbf{A}_s := \frac{\partial}{\partial r_N} \log \Theta_N^{-\frac{1}{2}}(\hat{X}_s(x)) ds, \quad d\mathbf{L}_s := d\mathbb{L}_s^{\text{Cut}(N)}(\hat{X}(x)). \quad (18)$$

Proof. To begin with, we see by the smooth coarea formula that

$$\begin{aligned} & \lim_{t \uparrow T} \int_M p_t^D(x, y) q_{T-t}(y, N) d\text{vol}_M(y) \\ &= \lim_{t \uparrow T} \int_N \int_{T_z N^\perp} (p_t^D(x, \exp) \mathbf{1}_{\mathcal{M}_z(N)} \theta_N)(\xi) (2\pi(T-t))^{-\frac{(m-n)}{2}} \exp \left[-\frac{\|\xi\|^2}{2(T-t)} \right] d\xi d\text{vol}_N(z) \\ &= \lim_{t \uparrow T} \int_N \int_{T_z N^\perp} (p_t^D(x, \exp) \mathbf{1}_{\mathcal{M}_z(N)} \theta_N)(\sqrt{T-t}\xi) (2\pi)^{-\frac{(m-n)}{2}} \exp \left[-\frac{\|\xi\|^2}{2} \right] d\xi d\text{vol}_N(z) \\ &= \int_N \int_{T_z N^\perp} (p_T^D(x, \exp) \mathbf{1}_{\mathcal{M}_z(N)} \theta_N)(0_z) (2\pi)^{-\frac{(m-n)}{2}} \exp \left[-\frac{\|\xi\|^2}{2} \right] d\xi d\text{vol}_N(z) \\ &= \int_N p_T^D(x, z) d\text{vol}_N(z) \end{aligned}$$

where 0_z denotes the origin of the vector space $T_z N^\perp$ and where the third equality follows from the compactness of the closure of D and the dominated convergence theorem. Then, denoting by $\{P_t^D : t \geq 0\}$ the Dirichlet heat semigroup, it follows from Girsanov's theorem that

$$\begin{aligned} \int_N p_T^D(x, z) d\text{vol}_N(z) &= \lim_{t \uparrow T} \int_M p_t^D(x, y) q_{T-t}(y, N) d\text{vol}_M(y) \\ &= \lim_{t \uparrow T} P_t^D q_{T-t}(\cdot, N)(x) \\ &= \lim_{t \uparrow T} \mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_D\}} q_{T-t}(\hat{X}_t(x), N) \hat{M}_t \right] \end{aligned}$$

with

$$\hat{M}_{t \wedge \hat{\tau}_D} = \exp \left[\int_0^{t \wedge \hat{\tau}_D} \frac{r_N(\hat{X}_s(x))}{T-s} \left\langle \frac{\partial}{\partial r_N}, \hat{U}_s dB_s \right\rangle - \frac{1}{2} \int_0^{t \wedge \hat{\tau}_D} \frac{r_N^2(\hat{X}_s(x))}{(T-s)^2} ds \right] \quad (19)$$

where \hat{U} is a horizontal lift of $\hat{X}(x)$ to the orthonormal frame bundle and where B is the associated \mathbb{R}^m -valued Brownian motion given by the antidevelopment of \hat{U} . Now, using

Itô's formula and formula (13) we deduce that

$$\begin{aligned}
& \log q_{T-(t \wedge \hat{\tau}_D)}(\hat{X}_{t \wedge \hat{\tau}_D}(x), N) \\
&= \log q_T(x, N) - \int_0^{t \wedge \hat{\tau}_D} \frac{r_N(\hat{X}_s(x))}{T-s} \left\langle \frac{\partial}{\partial r_N}, \hat{U}_s dB_s \right\rangle \\
&\quad + \int_0^{t \wedge \hat{\tau}_D} \frac{\partial}{\partial s} \log q_{T-s}(\hat{X}_s(x), N) ds + \int_0^{t \wedge \hat{\tau}_D} \frac{r_N^2(\hat{X}_s(x))}{(T-s)^2} ds ds \\
&\quad + \frac{1}{2} \int_0^{t \wedge \hat{\tau}_D} \Delta \log q_{T-s}(\hat{X}_s(x), N) ds \\
&\quad + \int_0^{t \wedge \hat{\tau}_D} \frac{r_N(\hat{X}_s(x))}{T-s} d\mathbb{L}_s^{\text{Cut}(N)}(\hat{X}(x)).
\end{aligned}$$

and so we can eliminate the stochastic integral in (19) by rearrangement and substitution. Finally, using the fact that

$$\frac{\partial}{\partial s} \log q_{T-s}(\cdot, N) = \frac{m-n}{2(T-s)} - \frac{r_N^2(\cdot)}{2(T-s)^2}$$

and also that

$$\Delta \log q_{T-s}(\cdot, N) = -\frac{\Delta r_N^2(\cdot)}{2(T-s)}$$

on $M \setminus \text{Cut}(N)$ together with formula (17) we can further simplify the resulting expression so as to obtain the desired formula. \square

Theorem 4.2. *Suppose that $\{D_i\}_{i=1}^\infty$ is an exhaustion of M by regular domains. Then, with $p_T^M(x, N)$ defined by (1), we have*

$$p_T^M(x, N) = q_T(x, N) \lim_{i \uparrow \infty} \lim_{t \uparrow T} \mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \exp \left[\int_0^t \frac{r_N(\hat{X}_s(x))}{T-s} (d\mathbf{A}_s + d\mathbf{L}_s) \right] \right]$$

where $d\mathbf{A}$ and $d\mathbf{L}$ are defined by (18).

Proof. Recalling that p^M is given as the limit of the increasing sequence of Dirichlet heat kernels p^{D_i} , it follows from the monotone convergence theorem that

$$p_T^M(x, N) = \lim_{i \uparrow \infty} p_T^{D_i}(x, N)$$

and so the result follows from Theorem 4.1. \square

The absolutely continuous random measure $d\mathbf{A}$ can be understood precisely in terms of Jacobi fields, using Heintze and Karcher's comparison theorem for $\frac{\partial}{\partial r_N} \log \theta_N$ which was carefully explained in [20, Section 3]. The singular random measure $d\mathbf{L}$ vanishes if the cut locus is polar for $\hat{X}(x)$. In particular, we have the following corollary of Theorem 4.2, which shows that we recover identity (6) in the Euclidean setting.

Corollary 4.3. *Suppose that M is stochastically complete, that the cut locus of N has Hausdorff dimension at most $m-2$ and that one of the following conditions is satisfied:*

- (A1)** $n \in \{0, \dots, m-1\}$, the sectional curvature of planes containing the radial direction vanishes and N is totally geodesic;
- (A2)** $n \in \{0, m-1\}$, the Ricci curvature in the radial direction vanishes and N is minimal.

Then we have

$$p_T^M(x, N) = q_T(x, N)$$

for all $x \in M$ and $T > 0$.

Proof. The corollary follows from Theorem 4.2, the Heintze-Karcher theorem and the monotone convergence theorem. \square

If N is a point then Theorem 4.2 provides a formula for the heat kernel itself. In this case, by the Heintze-Karcher theorem, the random measure $d\mathbf{A}$ has the property of being non-negative if the Ricci curvature of M is non-negative and non-positive if the Ricci curvature of M is non-positive. The Heintze-Karcher also yields a comparison theorem, stated below, in which we view \mathbb{H}_κ^n as a totally geodesic embedded submanifold of \mathbb{H}_κ^m and set $\mathbb{H}_0^n := \mathbb{R}^n$ and $\mathbb{H}_0^m := \mathbb{R}^m$. It is a partial generalization of the heat kernel comparison theorem of Cheeger and Yau [6].

Theorem 4.4. *Suppose that M is stochastically complete and that one of the following conditions is satisfied:*

- (B1)** $n \in \{0, \dots, m-1\}$, *the sectional curvature of planes containing the radial direction is bounded below by $\kappa \leq 0$ and N is totally geodesic;*
- (B2)** $n \in \{0, m-1\}$, *the Ricci curvature in the radial direction is bounded below by $(m-1)\kappa \leq 0$ and N is minimal.*

Then we have

$$p_T^M(x, N) \geq p_{T_\kappa}^{\mathbb{H}_\kappa^m}(y, \mathbb{H}_\kappa^n)$$

for any $y \in \mathbb{H}_\kappa^m$ satisfying $r_N(x) \leq r_{\mathbb{H}_\kappa^n}(y)$.

Proof. First suppose $\kappa = 0$. Then $\frac{\partial}{\partial r_N} \log \Theta_N \leq 0$, by [46, Theorem 1], so the assertion follows from Theorem 4.2 since \mathbf{L} is non-decreasing. So assume $\kappa < 0$ and define a function g_κ by

$$g_\kappa(t) := (m - n - 1) (\sqrt{-\kappa t} \coth(\sqrt{-\kappa t}) - 1) + n \sqrt{-\kappa t} \tanh(\sqrt{-\kappa t}).$$

Then, by the Heintze-Karcher theorem, it follows that

$$r_N \frac{\partial}{\partial r_N} \log \Theta_N \leq g_\kappa(r_N).$$

Since \mathbf{L} is non-decreasing with g_κ non-negative, it follows from this, Theorem 4.2 and dominated convergence that

$$p_T^M(x, N) \geq q_T(x, N) \mathbb{E} \left[\exp \left[- \int_0^T \frac{g_\kappa(r_N(\hat{X}_s(x)))}{2(T-s)} ds \right] \right].$$

Now let $\hat{Y}(y)$ be a Fermi bridge between y and \mathbb{H}_κ^n in time T such that its radial part $r_{\mathbb{H}_\kappa^n}(\hat{Y}(y))$ satisfies a stochastic differential equation driven by the one-dimensional Brownian motion β which drives the equation for $r_N(\hat{X}(x))$. Then the Heintze-Karcher theorem, applied to the coefficients of the stochastic differential equation for $r_N^2(\hat{X}(x))$, combined with the comparison theorem for solutions to stochastic differential equations given in [36], implies that $r_N(\hat{X}(x)) \leq r_{\mathbb{H}_\kappa^n}(\hat{Y}(y))$, almost surely. Since g_κ is non-decreasing this yields

$$p_T^M(x, N) \geq q_T(y, \mathbb{H}_\kappa^n) \mathbb{E} \left[\exp \left[- \int_0^T \frac{g_\kappa(r_{\mathbb{H}_\kappa^n}(\hat{Y}_s(y)))}{2(T-s)} ds \right] \right]. \quad (20)$$

But the right-hand side of (20) is equal to $p_T^{\mathbb{H}_\kappa^m}(y, \mathbb{H}_\kappa^n)$, by Theorem 4.2 and the fact that

$$g_\kappa(r_{\mathbb{H}_\kappa^n}) = r_{\mathbb{H}_\kappa^n} \frac{\partial}{\partial r_{\mathbb{H}_\kappa^n}} \log \Theta_{\mathbb{H}_\kappa^n},$$

so the result follows. \square

4.1 Hypersurface Local Time

If $X(x)$ is a Brownian motion starting at $x \in M$ then $r_N(X(x))$ is a continuous semi-martingale. For the case in which N is a hypersurface, the local time of this process at zero is denoted by $L^N(X(x))$ and referred to as the local time of $X(x)$ on N . In certain circumstances, such as when M is compact, it was explained in [46] that there is the formula

$$\mathbb{E}[L_t^N(X(x))] = \int_0^t p_s^M(x, N) ds. \quad (21)$$

In the hypersurface case, the integrated heat kernel is therefore given by the time derivative of a mean local time.

Example 4.5. *Viewing the 2-dimensional hyperbolic space \mathbb{H}^2 as an embedded totally geodesic submanifold of \mathbb{H}^3 , if $X(x)$ is a Brownian motion in \mathbb{H}^3 starting at x then $r_{\mathbb{H}^2}(X(x))$ is a Markov process and so, by results in [15] and a change of variables, formula (21) holds in this setting. Therefore, by (7), it follows that*

$$\begin{aligned} \lim_{t \uparrow \infty} \mathbb{E}[L_t^{\mathbb{H}^2}(X(x))] &= \operatorname{sech}(r_{\mathbb{H}^2}(x)) \int_0^\infty (2\pi t)^{-\frac{1}{2}} \exp\left[-\frac{r_{\mathbb{H}^2}^2(x)}{2t} - \frac{t}{2}\right] dt \\ &= \operatorname{sech}(r_{\mathbb{H}^2}(x)) \exp[-r_{\mathbb{H}^2}(x)]. \end{aligned}$$

In contrast, if \mathbb{R}^2 is viewed as a linear subspace of \mathbb{R}^3 with $X(x)$ a Brownian motion in \mathbb{R}^3 starting at x then

$$\lim_{t \uparrow \infty} \mathbb{E}[L_t^{\mathbb{R}^2}(X(x))] = \infty.$$

Theorem 4.4 yields the following comparison for the mean local time, where we retain the notational convention outlined before Theorem 4.4.

Corollary 4.6. *Suppose that N is a minimal hypersurface and that the Ricci curvature in the radial direction is bounded below by $(m-1)\kappa \leq 0$. Suppose also that $X(x)$ is a non-explosive Brownian motion on M starting at x , denote by $L^N(X(x))$ the local time on $X(x)$ on N and assume that formula (21) holds. Denote by $Y(y)$ a Brownian motion on \mathbb{H}_κ^m starting at y with $r_N(x) \leq r_{\mathbb{H}_\kappa^{m-1}}(y)$. Then*

$$\mathbb{E}[L_t^N(X(x))] \geq \mathbb{E}[L_t^{\mathbb{H}_\kappa^{m-1}}(Y(y))] \quad (22)$$

for all $t \geq 0$.

Proof. This follows directly from Theorem 4.4 and formula (21). \square

In particular, if $\kappa = 0$ with $x \in N$ and $r_{\mathbb{H}_0^{m-1}}(y) = 0$ then the right-hand side of (22) is equal to $\sqrt{\frac{2t}{\pi}}$.

5 Lower Bound and Asymptotic Relation

The heat kernel lower bounds of Cheeger and Yau [6] were proved using a bound on the Ricci curvature in the radial direction and a Laplacian comparison theorem. These are similar to the objects we use, although our method is quite different. Our method is closer in spirit to that of Wang [49], who also used stochastic techniques with unbounded curvature, but only in the one point case. Note that the constants C_1, C_2 and Λ appearing in the following theorem typically depend upon N .

Theorem 5.1. *Suppose that M is stochastically complete and that there exist constants $C_1, C_2 \geq 0$ such that one of the following conditions is satisfied off the union of N and its cut locus:*

- (C1)** *the sectional curvatures of planes containing the radial direction are bounded below by $-(C_1 + C_2 r_N)^2$ and there exists a constant $\Lambda \geq 0$ such that the principal curvatures of N are bounded in modulus by Λ ;*
- (C2)** *$n = 0$ and the Ricci curvature in the radial direction is bounded below by $-(m - 1)(C_1 + C_2 r_N)^2$;*
- (C3)** *$n = m - 1$ and the Ricci curvature in the radial direction is bounded below by $-(m - 1)(C_1 + C_2 r_N)^2$ and there exists a constant $\Lambda \geq 0$ such that the mean curvature of N is bounded in modulus by Λ .*

Then for each $T > 0$ there exists a constant $C \geq 0$, depending only on T, C_1, C_2, Λ, m and n , such that

$$p_t^M(x, N) \geq t^{-\frac{(m-n)}{2}} \exp \left[-\frac{r_N^2(x)}{2t} - C(1 + r_N^2(x)) \right]$$

for all $x \in M$ and $t \in (0, T]$.

Proof. By [46, Theorem 1], the curvature conditions imply

$$\frac{\partial}{\partial r_N} \log \Theta_N \leq \alpha + \beta r_N \tag{23}$$

with $\alpha := n\Lambda + (m - 1)C_1$ and $\beta := (m - 1)C_2$. Using the process $\hat{X}(x)$ constructed in Subsection 3.2 it follows from this, Theorem 4.2 and the fact that $\mathbb{L}^{\text{Cut}(N)}(\hat{X}(x))$ is non-decreasing that

$$p_T^M(x, N) \geq q_T(x, N) \lim_{i \uparrow \infty} \lim_{t \uparrow T} \mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \exp \left[-\int_0^t \frac{f(\hat{X}_s(x))}{T-s} ds \right] \right]$$

where $f := \frac{1}{2}(\alpha r_N + \beta r_N^2)$. For $t \in [0, T)$ we see that

$$\begin{aligned} \mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \exp \left[-\int_0^t \frac{f(\hat{X}_s(x))}{T-s} ds \right] &= \mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \sum_{p=0}^{\infty} \frac{\left(-\int_0^t \frac{f(\hat{X}_s(x))}{T-s} ds \right)^p}{p!} \\ &= \mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} - 1 + \sum_{p=0}^{\infty} \frac{\left(-\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \int_0^t \frac{f(\hat{X}_s(x))}{T-s} ds \right)^p}{p!} \\ &= \mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} - 1 + \exp \left[-\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \int_0^t \frac{f(\hat{X}_s(x))}{T-s} ds \right] \end{aligned}$$

from which it follows, by Jensen's inequality, that

$$\begin{aligned} & \mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \exp \left[- \int_0^t \frac{f(\hat{X}_s(x))}{T-s} ds \right] \right] \\ & \geq \mathbb{Q}_t \{t < \hat{\tau}_{D_i}\} - 1 + \exp \left[- \mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \int_0^t \frac{f(\hat{X}_s(x))}{T-s} ds \right] \right]. \end{aligned}$$

Inequality (23) implies by formula (17) that inequality (15) holds on $M \setminus \text{Cut}(N)$ with $\nu = m - n + \frac{\alpha}{2}$ and $\lambda = \frac{\alpha}{2} + \beta$ and so, by Theorem 3.1, we see that

$$\begin{aligned} \mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \int_0^t \frac{r_N^2(\hat{X}_s(x))}{T-s} ds \right] & \leq \int_0^t \frac{\mathbb{E}[\mathbf{1}_{\{s < \hat{\tau}_{D_i}\}} r_N^2(\hat{X}_s(x))]}{T-s} ds \\ & \leq (r_N^2(x) + \nu T) e^{\lambda T} \end{aligned}$$

for all $t \in [0, T)$. Similarly, by Corollary 3.2, we have

$$\begin{aligned} \mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \int_0^t \frac{r_N(\hat{X}_s(x))}{T-s} ds \right] & \leq \int_0^t \frac{\mathbb{E}[\mathbf{1}_{\{s < \hat{\tau}_{D_i}\}} r_N(\hat{X}_s(x))]}{T-s} ds \\ & \leq \left(\frac{r_N^2(x) + \nu t}{T} \right)^{\frac{1}{2}} e^{\frac{\lambda t}{2}} \int_0^t (T-s)^{-\frac{1}{2}} ds \\ & = 2 (r_N^2(x) + \nu t)^{\frac{1}{2}} e^{\frac{\lambda t}{2}} \left(1 - \left(\frac{T-t}{T} \right)^{\frac{1}{2}} \right) \\ & \leq 2 (r_N^2(x) + \nu T)^{\frac{1}{2}} e^{\frac{\lambda T}{2}} \end{aligned}$$

for all $t \in [0, T)$. Consequently

$$\mathbb{E} \left[\mathbf{1}_{\{t < \hat{\tau}_{D_i}\}} \int_0^t \frac{f(\hat{X}_s)}{T-s} ds \right] \leq \alpha (r_N^2(x) + \nu T)^{\frac{1}{2}} e^{\frac{\lambda T}{2}} + \frac{\beta}{2} (r_N^2(x) + \nu T) e^{\lambda T}$$

for all $t \in [0, T)$. Furthermore

$$\lim_{i \uparrow \infty} \lim_{t \uparrow T} \mathbb{Q}_t \{t < \hat{\tau}_{D_i}\} = 1,$$

by the bridge property and the monotone convergence theorem, so we have a lower bound

$$\begin{aligned} p_T^M(x, N) & \geq q_T(x, N) \exp \left[-\alpha (r_N^2(x) + \nu T)^{\frac{1}{2}} e^{\frac{\lambda T}{2}} - \frac{\beta}{2} (r_N^2(x) + \nu T) e^{\lambda T} \right] \\ & = q_T(x, N) \exp \left[-(n\Lambda + (m-1)C_1) (r_N^2(x) + \nu T)^{\frac{1}{2}} e^{\frac{\lambda T}{2}} \right. \\ & \quad \left. - \frac{(m-1)C_2}{2} (r_N^2(x) + \nu T) e^{\lambda T} \right] \end{aligned}$$

for all $T > 0$ and $x \in M$, from which the result follows. \square

Since the heat kernel is a positive fundamental solution to the heat equation, fixing one of the spatial variables for small times results in densities whose mass is localized around that fixed point. Riemannian manifolds are locally Euclidean, so for these small times the resulting densities should be comparable to the Gauss-Weierstrass kernel (2). The precise sense in which this is true is given by Varadhan's asymptotic relation,

originally proved in [47, 48], which states for the minimal heat kernel p^M of a complete Riemannian manifold M that

$$\lim_{t \downarrow 0} t \log p_t^M(x, y) = -\frac{d^2(x, y)}{2}$$

uniformly on compact subsets of $M \times M$. Similarly, the embedding of N in M is locally diffeomorphic to an affine embedding of \mathbb{R}^n in \mathbb{R}^m so for small times the integrated heat kernel $p^M(\cdot, N)$ should be comparable to the kernel $q(\cdot, N)$ defined by equation (5).

Theorem 5.2. *Suppose that M is a complete and connected Riemannian manifold of dimension m and that N is a compactly embedded submanifold of M of dimension $n \in \{0, \dots, m-1\}$. Then*

$$\lim_{t \downarrow 0} t \log p_t^M(x, N) = -\frac{d^2(x, N)}{2} \quad (24)$$

uniformly on compact subsets of M .

Proof. It is a simple matter to show that the left-hand side of (24) is less than or equal to the right-hand side, using Varadhan's relation and the fact that $r_N(x) \leq r_y(x)$ for all $y \in N$. To prove the opposite inequality assume first that M is compact. Then the result follows immediately from Theorem 5.1. So assume that M is non-compact, let K be any compact subset of M and for $x \in K$ and $y \in N$ denote by $\Gamma_{x,y}$ the set of all length-minimizing geodesic segments between x and y , viewed as a subset of M . Then $\Gamma_{x,y}$ contains (the image of) at least one such geodesic and by the triangle inequality the set

$$\Gamma_{K,N} := \bigcup_{x \in K, y \in N} \Gamma_{x,y}$$

is a bounded subset of M . Now let D be any regular domain in M containing $\Gamma_{K,N}$. Modify M outside of D so as to obtain a compact Riemannian manifold M_D (by doubling, for example) and suppose that D is sufficiently large so that

$$\lim_{t \downarrow 0} \frac{p_t^D(x, y)}{p_t^{M_D}(x, y)} = 1$$

uniformly for $x \in K$ and $y \in N$. This is the principle of *not feeling the boundary* considered by Hsu in [22]. Such D can always be found, as explained by Norris in [35], since we are assuming that M is non-compact. Then for all $\epsilon > 0$ there exists $t_{\epsilon,K} > 0$ such that for $t \in (0, t_{\epsilon,K})$ we have

$$(1 - \epsilon)p_t^{M_D}(x, N) \leq p_t^D(x, N) \leq p_t^M(x, N)$$

for all $x \in K$. It follows from this and the result in the compact case that

$$\lim_{t \downarrow 0} t \log p_t^M(x, N) \geq -\frac{d_{M_D}^2(x, N)}{2}$$

where d_{M_D} denotes the distance function on M_D . But since $\Gamma_{K,N}$ is contained in D it follows that $x \in K$ and $y \in N$ implies $d_{M_D}(x, y) \leq d(x, y)$. Therefore $d_{M_D}(x, N) \leq d(x, N)$ and the result follows. \square

Example 5.3. *Suppose that D is a regular domain. Then ∂D is a compactly embedded hypersurface and, according to [24, Theorem 5.2.6], one has*

$$\lim_{t \downarrow 0} t \log \mathbb{P}\{\tau_D(x) < t\} = -\frac{d^2(x, \partial D)}{2}$$

for all $x \in D$, where $\tau_D(x)$ denotes the first exit time from D of a Brownian motion $X(x)$ starting at x . On the other hand, according to [35, Theorem 1.2], one has

$$\lim_{t \downarrow 0} t \log p_t^M(x, D, x) = -\frac{d^2(x, \partial D)}{2}$$

for all $x \notin D$, where $p_t^M(x, D, y) := p_t^M(x, y) - p_t^{M \setminus \overline{D}}(x, y)$ is the measure of heat passing through D . Theorem 5.2 and formula (21) imply that, in either case, we have

$$\lim_{t \downarrow 0} t \log \frac{d}{dt} \mathbb{E} [L_t^{\partial D}(X(x))] = -\frac{d^2(x, \partial D)}{2}.$$

Example 5.4. If TM is equipped with the Sasaki metric [38, 39] then Theorem 5.2 implies

$$\lim_{t \downarrow 0} t \log p_t^{TM}(\xi, M) = -\frac{\|\xi\|^2}{2}$$

uniformly for ξ in compact subsets of TM .

One could also prove Theorem 5.2 using Ndumu's asymptotic expansion [32]. While this expansion is only valid away from the cut locus, it could be used in place of our lower bounds by connecting points in K to N with smooth curves, covering them with small balls and invoking the Markov property.

Alternatively, it was proved by Hino and Ramírez in [21], in the context of Dirichlet spaces, that

$$\lim_{t \downarrow 0} t \log \mathbb{P}\{X_0 \in A; X_t \in B\} = -\frac{d^2(A, B)}{2}$$

for all measurable sets A and B of positive measure, where X denotes the Markov process associated with the underlying local regular Dirichlet form and where d is an associated intrinsic distance. Using Sturm's upper bounds [44] and pointwise lower bounds, one can deduce from this the pointwise relation of Varadhan, as shown by Ramírez [37, Theorem 4.1]. A modification of this approach, replacing balls with tubular neighbourhoods and the pointwise lower bounds with our integrated heat kernel lower bounds, might also be used to deduce a relation similar to (24).

5.1 Obtaining an Upper Bound from the Lower Bound

Grigor'yan Hu and Lau showed in [17] how Gaussian upper bounds on the heat kernel can be obtained from lower bounds. While lower bounds are frequently obtained from upper bounds, as in the method of Aronson [2], this was the first result to go in the other direction. Having proved a lower bound on the integrated heat kernel, and since deducing an upper bound directly from Theorem 4.1 seems hard, it is natural for us to use this approach to obtain an upper bound. To do so, we require the following definition.

Definition 5.5. We say that vol_M is lower regular if there exist constants $T_0, C > 0$ such that

$$V_r(x) \geq Cr^m$$

for all $x \in M$ and $0 < r < \sqrt{T_0}$.

For example, if the injectivity radius of M is positive and the Ricci curvature is bounded above by a constant then vol_M is lower regular. For geometric conditions which imply a positive injectivity radius, see [5]. Now assume that vol_M is lower regular

and that the Ricci curvature of M is bounded below, by a constant. Then the lower bound of Theorem 5.1 implies a *near-diagonal lower estimate* of the form

$$p_t^M(x, y) \geq C' t^{-\frac{m}{2}}$$

for all $0 < t < T_0$ and $x, y \in M$ satisfying $d(x, y) < t^{\frac{1}{2}}$. This implies, by [17, Corollary 3.5], that there exist constants $c, \sigma^2 > 0$ (which might depend upon T_0) such that

$$p_t^M(x, y) \leq c t^{-\frac{m}{2}} \exp \left[-\frac{d^2(x, y)}{\sigma^2 t} \right]$$

for all $t \in (0, T_0)$ and $x, y \in M$. We therefore have the following theorem, by observing that $y \in N$ implies $r_y(x) \geq r_N(x)$.

Theorem 5.6. *Suppose that the Ricci curvature of M is bounded below by a constant, that vol_M is lower regular and that N is compact. Then there exist constants $c, \sigma^2 > 0$ (which might depend upon T_0) such that*

$$p_t^M(x, N) \leq c t^{-\frac{m}{2}} \exp \left[-\frac{r_N^2(x)}{\sigma^2 t} \right]$$

for all $x \in M$ and $t \in (0, T_0)$.

6 Gradient and Hessian Estimates

We are now in a position to prove gradient and Hessian estimates. They are global and hence require assumptions on curvature. For comparison, recall the derivative estimates of Cheng, Li and Yau [7]. They proved, in particular, that if the injectivity radius of M is positive and there exist constants A_i which bound the i^{th} covariant derivatives of the curvature tensor then for all $T > 0$ there exist positive constants $\alpha(m)$ and $C(m, A_0, \dots, A_{l-1}, T)$ such that

$$\|\nabla^l p_t^M(\cdot, y)_x\| \leq C(m, A_0, \dots, A_{l-1}, T) t^{-\frac{(m+l)}{2}} \exp \left[-\frac{\alpha(m) d^2(x, y)}{t} \right]$$

for all $x, y \in M$ and $t \in [0, T]$. Our estimates will be on the *logarithmic* derivatives of the *integrated* heat kernel. They will be derived using a slight modification of the method used by Stroock in [43]. We will show precisely how the gradient estimate is derived, whereas for the Hessian estimate we will refer to [43] for the details. In particular, to prove the gradient estimate we will need Theorem 5.1 and a formula for the derivative of the heat semigroup, which we will now state.

Assume that the Ricci curvature of M is bounded below, denote by $X(x)$ a Brownian motion on M starting at x and denote by U a horizontal lift of $X(x)$ with antidevelopment B . Denote also by $\{Q_s : s \geq 0\}$ the solution the ordinary differential equation

$$\begin{cases} \dot{Q}_s = -\frac{1}{2} \text{Ric}_{U_s} Q_s \\ Q_0 = U_0^{-1} \end{cases} \quad (25)$$

where $\text{Ric}_{U_s} := U_s^{-1} \text{Ric}^\sharp U_s$. Then for any bounded measurable function $f : M \rightarrow \mathbb{R}$ there is the formula

$$d(P_t f)_x = \mathbb{E} \left[f(X_t(x)) \frac{1}{t} \int_0^t \langle Q_s, dB_s \rangle \right] \quad (26)$$

for all $t > 0$. This differentiation formula was proved by Elworthy and Li in [12] and by Li in [26], using basic stochastic calculus, while another, related, approach was given

by Thalmaier in [45]. It was proved originally by Bismut in [4] for the case in which M is compact. An analogous formula was used by Stroock, whose estimate is based upon the following lemma, proved as [43, Lemma 6.45] using Jensen's inequality.

Lemma 6.1. *Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and ϕ a non-negative measurable function on Ω with $\mathbb{E}[\phi] = 1$. If ϕ is a measurable function on Ω such that $\phi\psi$ is integrable, then $\mathbb{E}[\phi\psi] \leq \mathbb{E}[\phi \log \phi] + \log \mathbb{E}[e^\psi]$.*

Theorem 6.2. *Suppose that M is a complete and connected Riemannian manifold of dimension m and that N is a compactly embedded submanifold of M of dimension $n \in \{0, \dots, m-1\}$. Suppose that the injectivity radius of M is positive and that the curvature tensor is bounded. Then for all $T > 0$ there exists a constant $C > 0$ such that*

$$\|\nabla \log p_t^M(\cdot, N)_x\|^2 \leq C \left(\frac{1}{t} + \frac{n}{t} \log \frac{1}{t} + \frac{d^2(x, N)}{t^2} \right) \quad (27)$$

for all $x \in M$ and $t \in (0, T]$.

Proof. For a bounded measurable positive function f and $\gamma \in \mathbb{R}$ set

$$\phi := \frac{f(X_t(x))}{P_t f(x)}, \quad \psi := \gamma \int_0^t \langle \mathcal{Q}_s v, dB_s \rangle$$

and see, by Lemma 6.1 and formula (26), that

$$\gamma t \frac{d(P_t f)_x}{P_t f(x)} \leq h_t(x; f) + \log \mathbb{E} \left[\exp \left[\gamma \int_0^t \langle \mathcal{Q}_s, dB_s \rangle \right] \right]$$

where

$$h_t(x; f) := \mathbb{E} \left[\frac{f(X_t(x))}{P_t f(x)} \log \frac{f(X_t(x))}{P_t f(x)} \right]. \quad (28)$$

Furthermore, denoting by R the minimum of the Ricci curvature, equation (25) implies

$$\log \mathbb{E} \left[\exp \left[\gamma \int_0^t \langle \mathcal{Q}_s v, dB_s \rangle \right] \right] \leq \frac{\gamma^2}{2} \int_0^t e^{-Rs} ds$$

and so, after minimizing over γ , we deduce

$$\left| \frac{d(P_t f)_x}{P_t f(x)} \right| \leq \frac{1}{t} \left(2h_t(x; f) \int_0^t e^{-Rs} ds \right)^{\frac{1}{2}}.$$

Now we choose $f(\cdot) = p_t^M(\cdot, N)$. Then $P_t f(x) = p_{2t}^M(x, N)$, by Tonelli's theorem, and for all $z \in M$ it follows that

$$h_t(x; p_t^M(\cdot, N)) \leq \sup_{z \in M} \log \left(\frac{p_t^M(z, N)}{p_{2t}^M(x, N)} \right). \quad (29)$$

The curvature assumptions imply that there exist constants $C_1, C_2 \geq 0$ such that the sectional curvatures of planes containing the radial direction are bounded below by $-(C_1 + C_2 r_N)^2$ and so, by Theorem 5.1, there exists a constant $c_1 \geq 0$, depending only on T, C_1, C_2, m and n , such that

$$p_{2t}^M(x, N) \geq (2t)^{-\frac{(m-n)}{2}} \exp \left[-\frac{r_N^2(x)}{4t} - c_1(1 + r_N^2(x)) \right] \quad (30)$$

for all $x \in M$ and $t \in (0, T]$. The assumptions also imply, by Theorem 5.6 and the Chapman-Kolmogorov equation, that there exists a constant $c_2 > 0$ such that

$$p_t^M(z, N) \leq c_2 t^{-\frac{m}{2}} \quad (31)$$

for all $t \in (0, T]$ and $z \in M$. Substituting the estimates (30) and (31) in to (29) proves the theorem. \square

It follows that the gradient estimate (27) holds automatically if M is compact. Furthermore, we can now prove the semimartingale property promised in Subsection 2.3. As pointed out by Thalmaier, the semimartingale property should hold without any assumptions on curvature. Güneysu proved this for the one point case using local estimates on the heat kernel. We do not yet have such estimates for the integrated heat kernel and so, for the time being, we make do with the assumptions of Theorem 6.2.

Corollary 6.3. *Under the assumptions of Theorem 6.2, the coordinate process on the bridge space $L_{x;N,T}$ is a semimartingale with respect to the bridge measure $\mathbb{P}^{x;N,T}$.*

Proof. It suffices to control the singularity in the drift close to the terminal time. Since the distance function r_N is defined as an infimum over N , it follows from (12) and Lemma 2.2 that there exists $\epsilon, C_\epsilon > 0$ such that $\mathbb{E}^{x;N,T}[r_N^2(X_t)] \leq C_\epsilon(T-t)$ for all $t \in (T-\epsilon, T]$. Therefore, by Theorem 6.2, there exists $C > 0$ such that

$$\begin{aligned} & \mathbb{E}^{x;N,T} \left[\int_{T-\epsilon}^T \|\nabla \log p_{T-t}^M(X_t, N)\| dt \right] \\ & \leq \int_{T-\epsilon}^T \mathbb{E}^{x;N,T} [\|\nabla \log p_{T-t}^M(X_t, N)\|^2]^{\frac{1}{2}} dt \\ & \leq \sqrt{C} \int_{T-\epsilon}^T \left(\frac{1}{T-t} + \frac{n}{T-t} \log \frac{1}{T-t} + \frac{\mathbb{E}^{x;N,T}[r_N^2(X_t)]}{(T-t)^2} \right)^{\frac{1}{2}} dt \\ & \leq \sqrt{C} \int_{T-\epsilon}^T \left(\frac{1}{T-t} + \frac{n}{T-t} \log \frac{1}{T-t} + \frac{C_\epsilon}{T-t} \right)^{\frac{1}{2}} dt \\ & < \infty \end{aligned}$$

and the result follows. \square

Note that we could have derived our gradient estimate using assumptions slightly weaker than those of Theorem 6.2, using Ricci curvature, the sectional curvature of planes containing the radial direction and lower regularity of the volume measure. To derive our Hessian estimate we need only add to these assumptions some suitable control on the curvature two-form and the derivative of the Ricci tensor. For simplicity, however, we state our Hessian estimate, as we did the gradient estimate, in terms of injectivity radius and the full curvature tensor. This way, our estimates can be easily compared to that of Cheng, Li and Yau, stated above for the case in which the injectivity radius of M is positive.

Corollary 6.4. *In addition to the assumptions of Theorem 6.2 suppose also that the first covariant derivative of the curvature tensor is bounded. Then for all $T > 0$ there exists a constant $C > 0$ such that*

$$\|\text{Hess} \log p_t^M(\cdot, N)_x\| \leq C \left(\frac{1}{t} + \frac{n}{t} \log \frac{1}{t} + \frac{d^2(x, N)}{t^2} \right)$$

for all $x \in M$ and $t \in (0, T]$.

Proof. It was proved by Stroock in [43] that for any continuous positive function f there exists $C > 0$ such that

$$t \frac{\|\text{Hess}(P_t f)_x\|}{P_t f(x)} \leq C(1 + h_t(x; f))$$

for all $x \in M$ and $t \in (0, T]$ where $h_t(x; f)$ is defined by (28). Choosing $f = p_t^M(\cdot, N)$, using the lower bound (30) and the on-diagonal upper bound (31) we obtain the corollary,

by Theorem 6.2 and the fact that

$$\text{Hess} \log P_t f = \frac{\text{Hess} P_t f}{P_t f} - d \log P_t f \otimes d \log P_t f$$

for all $t > 0$. □

Note that the constant C appearing in this estimate depends only on the injectivity radius of M , the length of the time interval T , the dimensions m and n and the bounds on the curvature tensor and its derivative.

6.1 Further Applications

The results in this article, in particular the estimate on the Hessian, will lead to a study of the space of continuous paths which end on a submanifold, potentially shedding light on the relationship between the geometry of the path space, the intrinsic geometry of the ambient manifold and the extrinsic geometry of the submanifold, extending recent work of Naber [28] to the submanifold setting. A related application, which also concerns the extrinsic geometry of the submanifold, is the study of mean curvature flow (MCF). MCF typically concerns the codimension one $n = m - 1$ case, where the normal component of the velocity of a point on a hypersurface N is given by the mean curvature at that point. An important object here is the F -functional

$$F_{x,t}(N) := (2\pi t)^{\frac{1}{2}} p_t^M(x, N)$$

introduced by Huisken [25] (and for which Theorem 4.2 now provides a probabilistic formula). The monotonicity of this functional under MCF is the reason for its importance. The *entropy* $\lambda(N)$ of the hypersurface (see [8, Section 3]) is then defined by $\lambda(N) := \sup_{x,t} F_{x,t}(N)$. The entropy captures something of the extrinsic complexity of N and determines which singularities can form as a result of the flow (and by Theorem 5.1, the entropy is always greater than or equal to 1). According to Subsection 4.1, one can interpret the F -functional as the ratio

$$F_{x,t}(N) = \frac{\frac{d}{dt} \mathbb{E}[L_t^N(X(x))]}{\frac{d}{dt} \mathbb{E}[L_t^{\mathbb{R}^{m-1}}(X(0))]}.$$

Indeed, the rate at which the expected local time increases for a hypersurface with high mean curvature ought to be greater than that for, say, a minimal hypersurface. So MCF has an interpretation in terms of hypersurface local time which could lead to further applications of our results.

Less abstractly, note that any non-degenerate Itô diffusion in \mathbb{R}^m , constrained to arrive in a submanifold at a fixed positive time, such as on a level set of a function or on the boundary of a domain, is a process of the type considered in this article (any non-degenerate diffusion on a smooth manifold induces a Riemannian metric, via the principle symbol of its generator, with respect to which it is a Brownian motion with drift). Multi-dimensional Itô diffusions constrained by the value of an observable at a fixed future time could have applications in finance. Alternatively, consider the case of a particle diffusing inside a biological cell and which is known to exit the cell, via the membrane, at a particular time. Then the possible trajectories of the particle are distributed as those of the corresponding conditioned diffusion (and the metric could even be used to capture some of the internal structure of the cell).

Finally, a connection to statistics is given by the problem of manifold denoising (see [19]), in which data is drawn from sampling a probability distribution supported

around a submanifold of Euclidean space. The idea is then to reconstruct the submanifold (or derive certain topological properties of it, as in [34]). The integrated heat kernel $p_{\sigma^2}^{\mathbb{R}^m}(x, N)$ (normalized by the hypersurface volume $\text{vol}_N(N)$) yields a probability distribution corresponding to the case in which observations of the submanifold are made subject to a Gaussian error.

References

- [1] Pablo Angulo Ardoy and Luis Guijarro, *Cut and singular loci up to codimension 3*, Ann. Inst. Fourier (Grenoble) **61** (2011), no. 4, 1655–1681 (2012). MR2951748
- [2] D. G. Aronson, *Bounds for the fundamental solution of a parabolic equation*, Bull. Amer. Math. Soc. **73** (1967), 890–896. MR0217444 (36 #534)
- [3] Dennis Barden and Hui Ling Le, *Itô correction terms for the radial parts of semimartingales on manifolds*, Probab. Theory Related Fields **101** (1995), no. 1, 133–146. MR1314177 (96a:58211)
- [4] Jean-Michel Bismut, *Large deviations and the Malliavin calculus*, Progress in Mathematics, vol. 45, Birkhäuser Boston, Inc., Boston, MA, 1984. MR755001 (86f:58150)
- [5] Jeff Cheeger, Mikhail Gromov, and Michael Taylor, *Finite propagation speed, kernel estimates for functions of the Laplace operator, and the geometry of complete Riemannian manifolds*, J. Differential Geom. **17** (1982), no. 1, 15–53. MR658471 (84b:58109)
- [6] Jeff Cheeger and Shing Tung Yau, *A lower bound for the heat kernel*, Comm. Pure Appl. Math. **34** (1981), no. 4, 465–480. MR615626 (82i:58065)
- [7] Siu Yuen Cheng, Peter Li, and Shing Tung Yau, *On the upper estimate of the heat kernel of a complete Riemannian manifold*, Amer. J. Math. **103** (1981), no. 5, 1021–1063. MR630777 (83c:58083)
- [8] Tobias Holck Colding, William P. Minicozzi, II, and Erik Kjær Pedersen, *Mean curvature flow*, Bull. Amer. Math. Soc. (N.S.) **52** (2015), no. 2, 297–333. MR3312634
- [9] Michael Cranston, Wilfrid S. Kendall, and Peter March, *The radial part of Brownian motion. II. Its life and times on the cut locus*, Probab. Theory Related Fields **96** (1993), no. 3, 353–368. MR1231929 (94g:60154)
- [10] E. B. Davies, *Explicit constants for Gaussian upper bounds on heat kernels*, Amer. J. Math. **109** (1987), no. 2, 319–333. MR882426 (88g:58174)
- [11] Bruce K. Driver, *A Cameron-Martin type quasi-invariance theorem for pinned Brownian motion on a compact Riemannian manifold*, Trans. Amer. Math. Soc. **342** (1994), no. 1, 375–395. MR1154540 (94e:60072)
- [12] K. D. Elworthy and X.-M. Li, *Formulae for the derivatives of heat semigroups*, J. Funct. Anal. **125** (1994), no. 1, 252–286. MR1297021 (95j:60087)
- [13] K. D. Elworthy and A. Truman, *The diffusion equation and classical mechanics: an elementary formula*, Stochastic processes in quantum theory and statistical physics (Marseille, 1981), Lecture Notes in Phys., vol. 173, Springer, Berlin, 1982, pp. 136–146. MR729719 (85g:58090)
- [14] A. Engoulatov, *A universal bound on the gradient of logarithm of the heat kernel for manifolds with bounded Ricci curvature*, J. Funct. Anal. **238** (2006), no. 2, 518–529. MR2253731 (2008b:58032)
- [15] Pat Fitzsimmons, Jim Pitman, and Marc Yor, *Markovian bridges: construction, Palm interpretation, and splicing*, Seminar on Stochastic Processes, 1992 (Seattle, WA, 1992), Progr. Probab., vol. 33, Birkhäuser Boston, Boston, MA, 1993, pp. 101–134. MR1278079 (95i:60070)
- [16] Alfred Gray, *Tubes*, second ed., Progress in Mathematics, vol. 221, Birkhäuser Verlag, Basel, 2004, With a preface by Vicente Miquel. MR2024928 (2004j:53001)

- [17] Alexander Grigor'yan, Jiaxin Hu, and Ka-Sing Lau, *Obtaining upper bounds of heat kernels from lower bounds*, Comm. Pure Appl. Math. **61** (2008), no. 5, 639–660. MR2388658 (2009b:35152)
- [18] Alexander Grigor'yan and Masakazu Noguchi, *The heat kernel on hyperbolic space*, Bull. London Math. Soc. **30** (1998), no. 6, 643–650. MR1642767 (99h:58183)
- [19] Matthias Hein and Markus Maier, *Manifold denoising*, Advances in Neural Information Processing Systems 19 (P. B. Schölkopf, J. C. Platt, and T. Hoffman, eds.), MIT Press, 2007, pp. 561–568.
- [20] Ernst Heintze and Hermann Karcher, *A general comparison theorem with applications to volume estimates for submanifolds*, Ann. Sci. École Norm. Sup. (4) **11** (1978), no. 4, 451–470. MR533065 (80i:53026)
- [21] Masanori Hino and José A. Ramírez, *Small-time Gaussian behavior of symmetric diffusion semigroups*, Ann. Probab. **31** (2003), no. 3, 1254–1295. MR1988472 (2004f:60167)
- [22] Elton P. Hsu, *On the principle of not feeling the boundary for diffusion processes*, J. London Math. Soc. (2) **51** (1995), no. 2, 373–382. MR1325580 (96a:58210)
- [23] Elton P. Hsu, *Estimates of derivatives of the heat kernel on a compact Riemannian manifold*, Proc. Amer. Math. Soc. **127** (1999), no. 12, 3739–3744. MR1618694 (2000c:58047)
- [24] Elton P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, Providence, RI, 2002. MR1882015 (2003c:58026)
- [25] Gerhard Huisken, *Asymptotic behavior for singularities of the mean curvature flow*, J. Differential Geom. **31** (1990), no. 1, 285–299. MR1030675
- [26] X.-M. Li, *Stochastic flows on noncompact manifolds*, Ph.D. thesis, University of Warwick, 1992.
- [27] Carlo Mantegazza and Andrea Carlo Menzucchi, *Hamilton-Jacobi equations and distance functions on Riemannian manifolds*, Appl. Math. Optim. **47** (2003), no. 1, 1–25. MR1941909 (2003h:49049)
- [28] A. Naber. Characterizations of Bounded Ricci Curvature on Smooth and NonSmooth Spaces. arXiv:1306.6512, 2013.
- [29] Masayoshi Nagase, *Expressions of the heat kernels on spheres by elementary functions and their recurrence relations*, Saitama Math. J. **27** (2010), 25–34 (2011). MR2752561 (2012c:58044)
- [30] M. Ndumu, *Brownian motion and the heat kernel on riemannian manifolds*, Ph.D. thesis, University of Warwick, 1989.
- [31] Martin N. Ndumu, *An integral formula for the heat kernel of tubular neighborhoods of complete (connected) Riemannian manifolds*, Potential Anal. **5** (1996), no. 4, 311–356. MR1401071 (98b:58174)
- [32] Martin N. Ndumu, *Brownian motion and Riemannian geometry in the neighbourhood of a submanifold*, Potential Anal. **34** (2011), no. 4, 309–343. MR2786702 (2012g:58049)
- [33] Martin Ngu Ndumu, *The heat kernel formula in a geodesic chart and some applications to the eigenvalue problem of the 3-sphere*, Probab. Theory Related Fields **88** (1991), no. 3, 343–361. MR1100896 (93a:58177)
- [34] Partha Niyogi, Stephen Smale, and Shmuel Weinberger, *Finding the homology of submanifolds with high confidence from random samples*, Discrete Comput. Geom. **39** (2008), no. 1–3, 419–441. MR2383768
- [35] James R. Norris, *Heat kernel asymptotics and the distance function in Lipschitz Riemannian manifolds*, Acta Math. **179** (1997), no. 1, 79–103. MR1484769 (99d:58167)
- [36] Shige Peng and Xuehong Zhu, *Necessary and sufficient condition for comparison theorem of 1-dimensional stochastic differential equations*, Stochastic Process. Appl. **116** (2006), no. 3, 370–380. MR2199554
- [37] José A. Ramírez, *Short-time asymptotics in Dirichlet spaces*, Comm. Pure Appl. Math. **54** (2001), no. 3, 259–293. MR1809739 (2002f:60108)
- [38] Shigeo Sasaki, *On the differential geometry of tangent bundles of Riemannian manifolds*, Tôhoku Math. J. (2) **10** (1958), 338–354. MR0112152 (22 #3007)

- [39] Shigeo Sasaki, *On the differential geometry of tangent bundles of Riemannian manifolds. II*, Tôhoku Math. J. (2) **14** (1962), 146–155. MR0145456 (26 #2987)
- [40] Alessandro Savo, *A mean-value lemma and applications*, Bull. Soc. Math. France **129** (2001), no. 4, 505–542. MR1894148 (2003k:58038)
- [41] O. G. Smolyanov, H. v. Weizsäcker, and O. Wittich, *Brownian motion on a manifold as limit of stepwise conditioned standard Brownian motions*, Stochastic processes, physics and geometry: new interplays, II (Leipzig, 1999), CMS Conf. Proc., vol. 29, Amer. Math. Soc., Providence, RI, 2000, pp. 589–602. MR1803450 (2001k:58078)
- [42] Daniel W. Stroock, *An estimate on the Hessian of the heat kernel*, Itô's stochastic calculus and probability theory, Springer, Tokyo, 1996, pp. 355–371. MR1439536 (97m:58212)
- [43] Daniel W. Stroock, *An introduction to the analysis of paths on a Riemannian manifold*, Mathematical Surveys and Monographs, vol. 74, American Mathematical Society, Providence, RI, 2000. MR1715265 (2001m:60187)
- [44] Karl-Theodor Sturm, *Analysis on local Dirichlet spaces. II. Upper Gaussian estimates for the fundamental solutions of parabolic equations*, Osaka J. Math. **32** (1995), no. 2, 275–312. MR1355744 (97b:35003)
- [45] Anton Thalmaier, *On the differentiation of heat semigroups and Poisson integrals*, Stochastics Stochastics Rep. **61** (1997), no. 3-4, 297–321. MR1488139 (98h:58204)
- [46] James Thompson, *Brownian motion and the distance to a submanifold*, Potential Anal. 2016, doi:10.1007/s11118-016-9553-2
- [47] S. R. S. Varadhan, *Diffusion processes in a small time interval*, Comm. Pure Appl. Math. **20** (1967), 659–685. MR0217881 (36 #970)
- [48] S. R. S. Varadhan, *On the behavior of the fundamental solution of the heat equation with variable coefficients*, Comm. Pure Appl. Math. **20** (1967), 431–455. MR0208191 (34 #8001)
- [49] Feng-Yu Wang, *Sharp explicit lower bounds of heat kernels*, Ann. Probab. **25** (1997), no. 4, 1995–2006. MR1487443 (99e:58176)
- [50] K.D. Watling, *Formulae for solutions (possibly degenerate) diffusion equations exhibiting semi-classical and small time asymptotics*, Ph.D. thesis, University of Warwick, 1986.
- [51] Keith D. Watling, *Formulae for the heat kernel of an elliptic operator exhibiting small-time asymptotics*, Stochastic mechanics and stochastic processes (Swansea, 1986), Lecture Notes in Math., vol. 1325, Springer, Berlin, 1988, pp. 167–180. MR960168 (90j:58146)
- [52] Keith D. Watling, *Formulae for solutions to (possibly degenerate) diffusion equations exhibiting semi-classical asymptotics*, Stochastics and quantum mechanics (Swansea, 1990), World Sci. Publ., River Edge, NJ, 1992, pp. 248–271. MR1182603 (93f:35107)